AD P002990

SELECTED FACTORS AFFECTING AIRCREW PERFORMANCE DURING SUSTAINED OPERATIONS

Kent A. Kimball, Ph.D.
Director, Biomedical Applications Research
US Army Aeromedical Research Laboratoy
P. O. Box 577
Fort Rucker, Alabama 36362

SUMMARY

Six recent graduates of initial rotary wing training flew a UH-1H helicopter for up to 4 hours while wearing each of three clothing ensembles. Each aviator wore the standard flight suit, the US chemical defense (CD) ensemble, and the United Kingdom (UK) CD ensemble in hot weather (mean WBGT 29°C). Skin temperatures (chest, thigh, upper arm, and calf), rectal temperature, heart rate, and preflight and postflight body weights were recorded. Cognitive testing was conducted preflight, postflight, and on non-flight days. Aviator performance measures were also obtained during flight.

Well acclimatized aviators were able to fly at least 2 hours without serious physiological impairment. Three of the six aviators terminated flight for medical reasons (heart rates >140 bpm or nausea) while wearing the US ensemble. In this study the susceptible subjects tended to be older and heavier. Heart rate was judged to be the most sensitive indicator of heat stress.

Cognitive testing and flight performance data obtained during this exercise did not demonstrate changes as a function of the type of flight ensemble worn during the test, nor did flight performance serve as a predictor of heat stress. Further investigations are required to verify the validity of these measures as indicators of heat loading in the operational setting.

INTRODUCTION

Intelligence experts assessing the source and strength of the enemy threat have postulated that the initial battle of the next war will be mid-intensity in scope, with enemy offensive operations directed at targets of opportunity in an attempt to seize and hold strategic natural, industrial, and military resources. These resources will then be utilized as a base from which to extend and increase the tempo of more massive continuous operations or as a settlement for a suspension of such hostilities. Enemy deployment will be directed against carefully chosen points, thus ensuring the potential for maximum penetration in these selected areas. The enemy will pursue the battle both day and night, permitting no periods for allied defense reinforcement, resupply, and reorganization. Weather and darkness will influence both enemy and allied activity during the battle, but their effect will be of diminishing importance as technological advances are applied to tactical problems and resultant hardware is fielded.

The projected duration of this first enemy surge is estimated to be from 1 to 3 days. That is, intelligence estimates are that the enemy will have achieved its initial objectives or its offensive activity will be neutralized within this time period after initial contact. A ready and immediately responsive force must be available on site to halt the enemy's advance, take the initiative, and oegin counteroffensive operations as required.

If allied forces are sufficient to halt enemy advances on at least selected fronts, the character of the war will change. Around-the-clock operations will continue, but the enemy will seek to utilize the mobility of its ground and air forces to assure maximum penetration at points which are considered to exhibit the least amount of defense activity. These operations will also be directed at rear echelon areas in an attempt to disrupt or destroy allied communications, command and control, and rearm and resupply centers which are vital for sustained soldier and weapons support. The character of these highly mobile operations will require a defense posture which permits rapid allied force massing at these points simultaneous with enemy buildup. Continuous operations will, of necessity, require the utmost from both men and materiel; and attrition, from not only hostile fire but task overload and fatigue, must be considered when determining defense strength requirements.

Current military doctrine calls for extensive use of the helicopter for support, mobility, and fire power to counter the enemy threat. Army aviation units will participate in combined arms operations by providing highly maneuverable anti-armor fire power and light artillery mobility, as well as troop resupply, communications, and medical evacuation.[1] These missions may need to be accomplished in spite of highly effective air defense artillery threat and/or under conditions where the enemy enjoys air superiority. To counter this threat, Army doctrine requires that rotary wing flight be conducted as close to the earth's surface as possible in the combat environment. It has also been recognized that, unlike previous combat situations where the majority of tactical flight was accomplished in daylight, a significant amount of night terrain flight will be required in order to effectively complete the aviation mission.

The scenaric which has been discussed to this point delineates a very fluid, rapidly changing, violent battlefield environment where maximally efficient use of combat resources is vital for both tactical and strategic success. Recent intelligence and published changes in Warsaw Pact Military Doctrine require significant additions to this threat scenario which serve to increase exponentially the problems for the combat aviator. Although the threat of both strategic and tactical weapons as well as chemical weapons has been considered in US military doctrine, the latter has not received emphasis until recently when public attention was focused on their deployment and use by the Soviets and their counterparts in Southeast Asia and Afghanistan.

In March 1982, then Secretary of State Alexander Haig provided a comprehensive report on this subject to the Congress. The following is an excerpt from this report.

> "Evidence accumulated since World War II clearly shows that the Soviets have been extensively involved in preparations for large-scale offensive and defensive chemical warfare. Chemical warfare agents and delivery systems developed by the Soviets have been identified, along with production and storage areas within the USSR and continuing research, development, and testing activities at the major Soviet chemical proving grounds. Soviet military forces are extensively equipped and trained for operations in a chemically contaminated environment. The Soviets have shown a strong interest in improving or enhancing their standard agents for greater reliability and effect. Their large chemical and biological development effort has led them to investigate other kinds of chemical warfare agents, particularly the toxins."[2]

Based on Soviet activity and doctrine, it is likely that US Army aviators will be exposed to chemical agents both in flight and on the ground during the next war.[3] On such a battlefield the aviator will be at serious risk. To counter this threat, aircrews will be required to wear chemical defense (CD) ensembles on a continuous basis during the performance of their mission. The ability of the pilot to effectively operate his aircraft while in a CD ensemble and sustain such performance is a critical element of mission success on the chemically contaminated battlefield.

Current CD ensembles impose physiological limitations such as restrictions to pulmonary function, impairment of thermoregulatory mechanisms, and physical limitations such as impaired vision, limited manual dexterity,[4] and distorted communications. The pulmonary function restriction is due to breathing resistance caused by the charcoal filter through which the aviator must breathe. The thermoregulatory impairment is caused by the increase in insulation around the man due to the thickness of the CD ensemble. Heat related problems vary in severity from the trivial to the deadly; i.e., heat cramps, heat exhaustion, heat pyrexia, and heat stroke. Visual restrictions are caused by present mask design; i.e., narrowed field of view, reduced peripheral vision, and distortion caused by optical faceplate characteristics.[5] Limited manual dexterity is due to the bulkiness of butyl rubber gloves which must be worn over standard Nomex flight gloves. This reduction in dexterity creates the potential for degraded performance of inflight emergency procedures such as resetting circuit breakers and/or moving switches while maintaining aircraft control. The ability to communicate both inside and outside the aircraft is critical for safety and mission accomplishment. The present US aviator's mask does not have a voice emitter; therefore, communications outside during preflight, refueling, and rearming are difficult.[6] The combination of these factors may substantially limit aviator performance and reduce efficient mission accomplishment.

Thus, in reviewing the potential battlefield environment and the aircrew requirements for survival and optimum mission accomplishment, tactical planners and field commanders are faced with a further dilemma. Succinctly, given the already substantial physiological cost of human staying power for sustained and continuous operations, what is the impact created by the further requirement to fight and survive on the chemically contaminated battlefield?

In an attempt to gain insight into the performance potential of aviators operating under the constraints of CD ensembles, the US Army Aeromedical Research Laboratory (USAARL) conducted a controlled flight exercise in the summer of 1981. Detailed descriptions of this research exercise are provided in USAARL Reports 82-9,[7] 83-4,[8] and 83-6.[9] The specific purpose was to assess pilot physiological, cognitive, and flight performance in a standard flight suit as compared with the current United Kingdom (UK) and the US CD ensembles when worn in a hot environment. The remainder of this paper will delineate the method and procedures utilized for the test and discuss the results obtained on selected factors of performance, physiology, and cognitive function.

METHOD

Six male volunteer aviators flew a JUH-1H helicopter while wearing each of three ensembles: the standard US Army Nomex flight suit (ST), the US Army aircrew chemical defense ensemble (US), or the United Kingdom aircrew chemical defense ensemble (UK). Each day the subjects wore a different suit/ensemble and the order in which the suits were flown was counterbalanced for order-of-wear effects. Subjects were asked to fly 4 hours per day on 3 alternating days. During each flight, they were asked to fly a series of maneuvers repetitively. The series consisted of a 50-foot out-of-ground-effect hover, a lateral hover exercise, and a precision flight maneuver. The series took about 40 minutes to complete and was repeated until the 4 hours were finished or the flight was terminated because the subject exceeded established heat safety criteria. These safety limits were: heart rate >140 beats per minute (bpm) for 15 minutes, core temperature >38.5°C or mean skin temperature within 0.5°C of core temperature.[10 11] Table 1 represents a listing of the physiological measures taken during the flight days. All data were recorded on magnetic tape via a Helicopter Inflight Monitoring System (HIMS) for later analysis.[12]

TABLE 1

PHYSIOLOGICAL PARAMETERS

- Temperature - Skin/Rectal/WBGT
- Body Weight - Before/After
- Heart Rate
- Body Fluids - Blood/Urine

The information on heart rate, skin and rectal temperature, and Wet Bulb Globe Temperature (WBGT) was checked routinely by a medical observer to assure compliance with safety criteria.

Each subject spent from Sunday evening through Friday afternoon at USAARL's remote research facility. All meals were prepared under the direction of a dietitian to minimize weight gain/loss. Blood samples were taken on Thursday prior to the flight week and on Friday afternoon after the last flight. Urine specimens were collected and stored for later analysis. Before and after flight and at the same time on rest days, each subject completed a psychomotor/cognitive test using a microcomputer. On flight days, each subject was weighed before and after his flight as were pieces of the ensemble so that the accumulation of sweat within the ensemble could be measured. After application of temperature and electrocardiograph (ECG) sensors, the subject was weighed. After donning the ensemble he was reweighed and his vital signs recorded. He then proceeded to the aircraft which was parked near the facility with rotor blades turning. In order to minimize initial heat loading, the subject was not required to perform preflight checks but immediately entered the aircraft and executed the flight maneuvers as instructed by the safety pilot.

After one hour of flight the subject was provided water to drink. The quantity of water consumed by the subject was recorded. At the 2-hour mark, the helicopter returned to the landing field for refueling. The subject exited the aircraft and walked to an area adjacent to the research facility where he sat in the shade and again drank water. After refueling, the subject resumed flying. After 4 hours, or sooner if the subject exceeded safety criteria, the subject returned to the research facility. Following the flight, the subject's vital signs were recorded and he was weighed before and after doffing the ensemble.

RESULTS AND DISCUSSION

Physiological Factors

Total time, as shown in Table 2, covers the time from donning the ensemble through the flight to doffing the ensemble. The mean total time across subjects for the ST suit was 4.66 hours, for the UK ensemble, 4.19 hours, and for the US ensemble, 4.00 hours.

TABLE 2

TOTAL TIME IN UNIFORM (HOURS)

SUBJECT	SUIT ST	US	UK
A	4.65	4.65	2.83*
B	4.43	4.65	4.28
C	4.58	2.93†	3.38
D	4.97	2.87	4.92
E	4.68	4.40	2.33†
F	§	3.42	3.12†
MEAN (SD)#	4.60 (0.20)	4.00 (0.81)	4.19 (0.77)

*Flight terminated prematurely due to application of overly conservative safety criteria relating to convergence of skin and rectal temperature.
†Flight shortened by weather.
§Flight not flown due to weather.
#Means do not include items marked * and †.

The second measure of exposure was flying time taken from the time of entering the aircraft until termination of the last flight, including water and refueling breaks. As shown in Table 3, mean flying times were shortest in the US ensemble (3.17 hours), longer in the UK ensemble (3.79 hours), and the longest in the ST suit (3.89 hours).

TABLE 3

FLYING TIME (HOURS)

SUBJECT	SUIT ST	US	UK
A	4.03	3.28*	2.12†
B	3.75	3.80	3.65
C	3.95	1.87§	3.73
D	3.83	2.25*	4.00
E	3.00#	3.75	1.75§
F	-§	2.75*	2.63§
Mean**	3.89	3.17	3.79
±SD	.12	.66	.18

*Terminated flights for exceeding medical safety limits or voluntarily while near safety limits.
†Terminated early due to overly conservative safety criteria relating to convergence of skin and rectal temperatures.
§Weather
#Precautionary landing.
**Means do not include items marked †, §, and #.

These means include two subjects terminated by the medical observer for heart rates exceeding 140 bpm and one subject who withdrew while wearing the US ensemble because he was weak and nauseous. This subject's mean heart rate was 136 bpm with peaks above 140 bpm at withdrawal. No other subject terminated flight early for medical reasons.

One of the consequences of flying in hot weather is the loss of large amounts of sweat as the body attempts to cool itself by convective means. Sweat loss becomes critically important if it results in appreciable dehydration.

Table 4 presents both the uncorrected weight loss, weight (before)-weight (after) which is indicative of the relative dehydration of the pilot and the corrected weight loss, weight (before)+water-urine-weight (after), which is indicative of sweat loss.

The mean corrected weight loss for the ST suit was 0.95 kg; for the US ensemble, 1.29 kg; and for the UK ensemble, 1.43 kg. As judged by the uncorrected weight loss, the state of dehydration was slight in the ST suit and the US ensemble while the UK ensemble showed a more substantial dehydration effect.

TABLE 4

WEIGHT LOSS (kg) CORRECTED FOR WATER INTAKE

	SUIT					
	ST		US		UK	
Subject	Uncorrected	Corrected	Uncorrected	Corrected	Uncorrected	Corrected
A	+.15	-.46	+1.26	+.10	-1.00	-1.81†
B	-.05	0.61	-.80	-2.18	-1.94	-2.70
C	-.72	-1.49	-.50†	-§	+.01	-.60
D	-.62	-1.49	-.17	-1.25	-.35	-1.00
E	-.30	-.71	-.51	-1.73	-.93	-1.26†
F	-	-†	-.39	-1.37	-.80	-1.12†
Mean	-.31	-.95	-.12	-1.29	-.81	-1.43
±SD	0.37	0.50	0.81	0.86	0.73	1.12

*Gain = (+), loss = (-).
†Not included in mean. Flights shortened or not flown (subject F) due to weather or aborted too early (subject A) due to overly conservative safety criteria relating to convergence of skin and rectal temperature. Corrected WT = WT (before) + water - urine - WT (after).
§Water consumption data incomplete.

TABLE 5

WATER CONSUMPTION (kg)

	SUIT		
Subject	ST	US	UK
A	0.608	1.358	0.808*
B	0.831	1.383	0.885
C	0.769	-†	0.614
D	0.866	1.078	0.647
E	0.413	1.221	0.332†
F	-†	0.982	0.315†
Mean§	0.70	1.20	0.72
±SD	0.19	0.17	0.15

*Subject terminated early due to overly conservative safety criteria relating to convergence of skin to rectal temperature.
†Missing or altered data due to weather.
§Data not included in means marked with * and †.

The water consumed ad libitum by the pilots is referenced in Table 5. On the average, pilots consumed equivalent amounts of water while wearing the standard flight suit and UK ensemble, but they consumed almost twice as much while wearing the US ensemble.

TABLE 6

WEIGHT CHANGE OF UNIFORM (kg)

Subject	SUIT		
	ST	US	UK
A	0.211	0.514	0.343*
B	(0.114)†	0.594	1.007
C	(0.073)	0.303*	(0.017)
D	(0.120)	1.293	0.229
E	0.086	1.192	0.498*
F	-*	0.511	0.438*
Mean	-0.002	0.821	0.406
±SD	0.145	0.388	0.535
Retained in Uniform§	0.2%	64%	28%

*Not included in Means. Flights shortened or not flown (subject F) due to weather or aborted too early (subject A) due to overly conservative safety criteria related to convergence of skin and rectal temperature.
† ( ) Weight loss during flight.
§ $\frac{\text{WT Change in Uniform}}{\text{Corrected WT Loss (Table 5)}} \times 100$

Sweat which is lost from the body will either evaporate, drip off, or be absorbed by the ensemble. Table 6 presents a summary of the weight changes in the uniforms, presumably attributable to the absorption of sweat. The ST suit lost an average of 0.002 kg while the US and UK ensembles gained 0.821 and 0.406 kg, respectively. When compared to the corrected weight loss, the ST suit retained essentially none of the available sweat while both the UK and US retained considerably more.

Since heat stress was a major concern in this exercise, the physiological measures concentrated on temperature changes and heart rate changes as a function of time. The suits were exposed to mean WGBTs of 28.92°C (ST), 29.20°C (US), and 29.01°C (UK). These WBGT readings were measured at the pilot station within the aircraft.

To illustrate the effects encountered during this exercise, consider Figures 1, 2, 3, and 4. Figures 1 and 2 show the rectal temperature and heart rate for one subject as a function of time while wearing each of the three ensembles. His core temperature remained stable in the ST and UK ensembles, while it showed a continuing rise after the refueling stop in the US ensemble. His heart rate (Figure 2) showed a similar pattern with mean heart rates for ST (83.4) and UK (98.4) generally stable with small fluctuations presumably driven by moment-to-moment stresses of flight, while the heart rate in the US ensemble showed a substantial rise subsequent to the refueling stop after 2 hours of flight. This rising heart rate resulted in the termination of one subject's flight after 3.28 hours for exceeding safety limits (HR>140bpm).

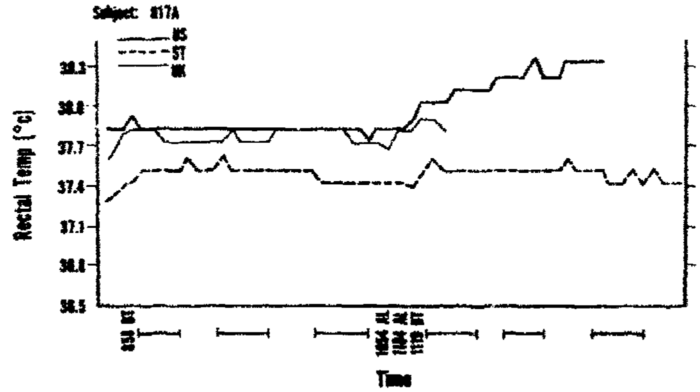

FIGURE 1. Plot of Rectal Temperature as a Function of Flight Time. BT = before takeoff, AL = after landing, |——| = time of precision flight maneuver, and blank = time of 50-foot hover and lateral hover exercise. ST = standard one-piece flight suit, US = US chemical defense ensemble, and UK = United Kingdom CD ensemble. Refueling took place between 1054 and 1119 hours. Data plotted every 5 minutes. Subject aborted flight in US for heart rate >140 bpm. He was terminated early for an overly conservative criterion of convergence of chest and rectal temperature while wearing UK.

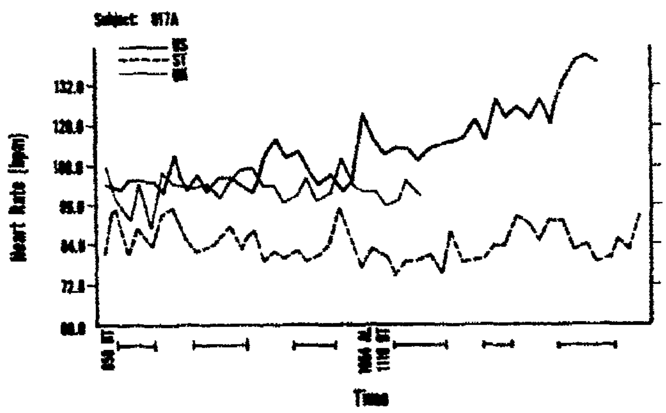

FIGURE 2. Heart Rate as a Function of Flight Time. Data plotted every 5 minutes except during refueling which occurred between 1054 and 1119 hours. Same flights as depicted in Figure 4.

By way of contrast, consider the same information from another subject who flew the same days (Figures 3 and 4). This second subject flew all three ensembles for the required period of time: ST=3.75, US=3.80, and UK=3.65 hours. His rectal temperature in the US ensemble rose after the refueling stop but, unlike the other subject, did not continue to rise during the subsequent flight. His corresponding heart rate did not rise so he was able to continue to fly.

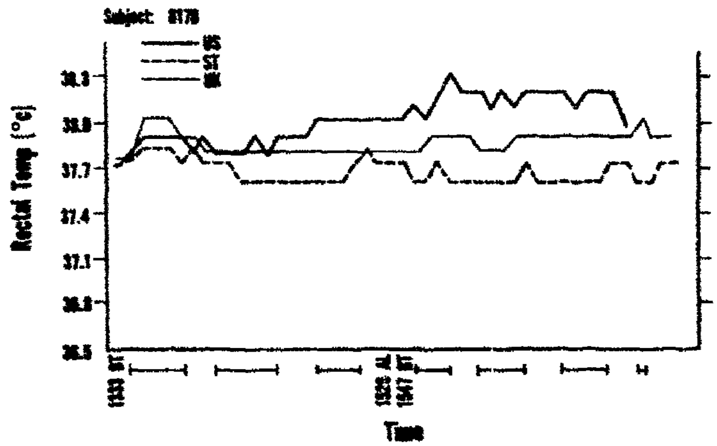

FIGURE 3. Rectal Temperatures as a Function of Flight Time. Refueling occurred between 1525 and 1547 hours. Subject 817B completed all flights.

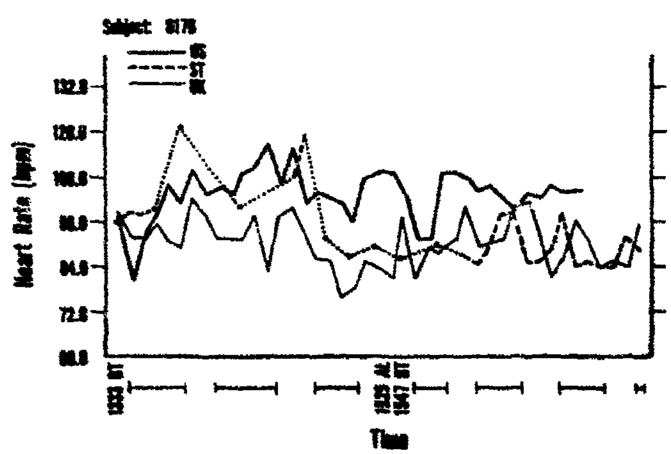

FIGURE 4. Heart Rate as a Function of Flight Time. Refueling took place between 1525 and 1547 hours. Dotted line indicates missing data due to noisy electrodes during flight with ST suit. All heart rates were well below safety cutoff of 140 bpm.

In an effort to follow the physiological behavior of the six subjects as a group, mean rectal temperatures are plotted across time in Figure 5. Data presented here shows that up to 2 hours in each ensemble produced only minor changes in rectal temperature. The US ensemble showed a stable mean core temperature; the UK tended to cool off; and the ST tended to warm up slightly. As mentioned previously, after the 2-hour refueling, a substantial rise in mean core temperature was measured for the US ensemble.

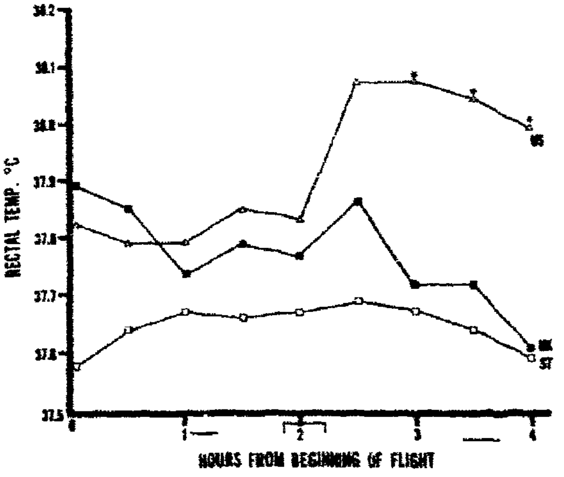

FIGURE 5. Rectal Temperature vs Time for Each Suit. ——— indicates a water break and ┌──┐ indicates a refueling and water break. * indicates drop out of subjects (US) from the initial 6 to 5 to 3 to 2. Beginning with the 2-hour point, all succeeding points on the UK curve represent data from three subjects except the last which represents data from two.

Perhaps a less direct, but more sensitive indicator of heat stress is heart rate. One way to look at the effect of flying in various ensembles on heart rate is to determine the change in rate from baseline and the mean heart rate during the last few minutes of flight. These changes in heart rate (Table 7) indicate that the standard uniform is least stressful with a mean increase of 9.3 bpm followed by the UK uniform with a mean increase of 27.8 bpm. The US ensemble exhibited a 48.5 bpm increase over baseline.

TABLE 7

CHANGE IN HEART RATE (HR) OVER BASELINE

		Change in HR		
			SUIT	
Subject	Baseline HR	ST	US	UK
A	80	3.6	49.8*	17.7†
B	75	8.4	21.2	13.2
C	73	-2.4	23.0†	25.4
D	70	19.7	71.1*	44.7
E	60	17.0	46.2	32.0†
F	76	†	54.0§	33.0†
Mean		9.3	48.5	27.8
±SD		9.2	18.0	15.9

*Terminated by medical observer (HR >140).
†Not included in means. Flights shortened or not flown (subject F) due to weather or aborted too early (subject A) due to overly conservative criterion regarding convergence of skin and rectal temperature.
§Terminated flight by self; subject had high HR, nausea, and weakness.

When tested for the usual spectrum of factors (SMA18), the pre- and poststudy blood samples showed no remarkable changes. Thus, the level of stress experienced by the subjects was transient and insufficient to cause major shifts in blood chemistry. Likewise, there was no shift in urine osmolality indicative of dehydration; although there was a tendency for pilots to void greater than normal volumes of urine containing excess sodium on rest days.

Taken together, these biochemical measures suggest that as a group, the pilots were not stressed to any great extent. One would expect, however, that had we asked the pilots to fly more demanding scenarios for longer periods of time, their physiological and biochemical parameters would have reflected the increased stress and workload.

Cognitive Factors

Psychological and psychomotor tests consisted of subtests selected from the performance Assessment Battery (PAB) developed by the Division of Neuropsychiatry, Walter Reed Army Institute of Research (WRAIR). These tests were administered to all subjects prior to suiting up for flight and as soon after flight as possible. The following subtests were used:

Mood Scale. Subjects were asked to rate their agreement to 65 mood descriptors (such as "anxious") on a 1 (none) to 5 (extreme) scale. Order of presentation was randomized for each test with some words repeated as controls.

Feeling/Tone.[13] Subjects were asked to rate their current level of fatigue by stating whether or not they felt "better than," "same as," or "worse than" the activity level descriptor.

Encode/Decode.[14] Subjects were given an arbitrary coding system which related letters to two-digit numbers and were asked to encode or decode purported map coordinates according to a set of simple rules. They were to do as many as possible in 2 minutes.

Target Recognition.[15] Subjects were given two target letters and asked to determine if both letters occurred in a list of 30 letters or if one or both letters did not occur in the list. They were to do as many as possible in 2 minutes.

Logical Reasoning.[16] Subjects were given a sentence which claimed to describe the order of the two letters (AB or BA) which followed the sentence. Their task was to determine if the order described was the same as that given. They were to complete as many as possible in 2 minutes.

Serial Math.[17] Subjects were asked to watch a briefly presented string of characters (approximately .25 second per character). The first two were numbers in the range of zero to nine with the third character being an add or subtract sign. The task was to perform the operation on the numbers and either add or subtract 10 from the result if the results met certain criteria. Subjects were to complete as many problems as possible in 2 minutes.

Reaction Time.[18] Subjects were presented with a four-choice reaction time task. This task presented the subject with four lights arranged in a square pattern. The subject's task was to determine which light was illuminated and press the button in the corresponding position as quickly as possible. The task was presented repetitively for 8 minutes.

Subtests of PAB were scored by computer for the number attempted, percent correct, reaction time to correct response (RTcor), and reaction time to an incorrect response (RTerr). Information on mood was converted into mean score in the categories of mood (good to bad), hostility (friendly to hostile), happiness (happy to unhappy) and depression (in-the-dumps to on-top-of-the-world). Since performance is susceptible to circadian changes[19] as well as to individual differences, direct comparison of the data is difficult to interpret. In order to control for these outside influences, the raw data were converted into percent of change from baseline (pretest) using the formula (A-B)/B where B was the pretest score and A the posttest score. In this manner, any one experimental manipulation was represented by a percent of change score which was the composite of the pretest and posttest scores. Positive scores indicate increases in posttest scores over pretest scores. Evaluation of this number required reference to either control or experimental data dependent upon the comparison desired.

Statistical significance was determined by means of a Randomized Block ANOVA with Replicates[20] The factors used were percent change from baseline on control days, ST suit days, UK CD ensemble days, and US CD ensemble days. None of the subtests of PAB exceeded the $p=.05$ value and it was concluded that there were no statistically significant effects associated with any of the factors. Similarly, self-report of mood failed to show any significant differences across factors.

The purport... stressor in this study was heat. Analysis of the physiological data revealed that subjects showed markedly different physiological responses to the experimental conditions. Therefore, experimental data were divided into three categories irrespective of suit and based soleiy upon physiological response. The three categories selected were slight, moderate, and severe heat stress. Placement into a category was determined by a physiologist who had no knowledge of the outcome of the psychological testing and was given only the category titles, "slight," "moderate," and "severe," without specific placement criteria. The convention adopted was that subjects withdrawn from an experimental condition because they exceeded heat and/or safety criteria would be judged as severely heat-stressed subjects, those with consistently elevated heart rates or temperatures but less than the heat safety criteria would be judged as moderately heat-stressed, and the remaining would be judged as slightly heat-stressed. Accordingly, three instances of severe heat stress, seven instances of moderate heat stress, and six instances of slight heat stress were identified. (Two flights were not flown because of inclement weather.)

The data for these groups as well as for the nonflight (control) days were averaged and are presented by subtest. These arbitrary groupings crossed the original group bounds and left three groups which were composed of partial replicates of unequal size, and generally violated most assumptions concerning population homogeneity. The results are, therefore, trends without statistical confirmation.

Table 8 presents the percent of change data for the logical reasoning test. Again, positive percentages indicate increased posttest scores relative to the pretest and negative percentages indicated decreased posttest scores relative to the pretest. The most notable features are the slight changes in accuracy (percent correct) and the large changes in reaction time to an incorrect response (RTerr) as opposed to the small changes in reaction time to correct response (RTcor).

TABLE 8

PERCENT CHANGE IN LOGICAL REASONING TESTS

	NUMBER ATTEMPTED	PERCENT CORRECT	RT CORRECT	RT ERROR
CONTROL	2.0	-2.0	1.0	-17.0
SLIGHT	5.0	4.0	0.0	20.0
MODERATE	3.0	3.0	-11.0	-9.0
SEVERE	2.0	1.0	-3.0	-13.0

RT = Reaction time

Similar data is presented in Table 9 for the target recognition task. The data indicate a greater degree of change in accuracy than in the logical reasoning test. RTerr changed by approximately the same magnitude as in the logical reasoning test but in the opposite direction, towards longer deliberations.

TABLE 9

PERCENT CHANGE IN TARGET RECOGNITION TESTS

	NUMBER ATTEMPTED	PERCENT CORRECT	RT CORRECT	RT ERROR
CONTROL	5.0	-1.0	-2.0	-29.0
SLIGHT	7.0	2.0	-9.0	-29.0
MODERATE	2.0	3.0	-7.0	*
SEVERE	-3.0	-5.0	-6.0	20.0

*Not computable
RT = Reaction time

Table 10 presents the results of the serial math test. Accuracy in this test was not as sensitive to the imposed heat stress as the previous tests. RTcor and RTerr both showed changes with the severely stressed group working correct answers faster while working incorrect answers slower.

TABLE 10

PERCENT CHANGE IN SERIAL MATH TESTS

	NUMBER ATTEMPTED	PERCENT CORRECT	RT CORRECT	RT ERROR
CONTROL	-2.0	-1.0	-5.0	51.0
SLIGHT	1.0	6.0	5.0	23.0
MODERATE	5.0	6.0	-7.0	10.0
SEVERE	10.0	4.0	-19.0	7.0

RT = Reaction time

The data for the encode/decode test are summarized in Table 11. Accuracy showed similar changes in this test as in the previous tests, but magnitude of change was not as great as in the serial math test. RTcor also demonstrated smaller changes than previously seen.

TABLE 11

PERCENT CHANGE IN ENCODE/DECODE TESTS

	NUMBER ATTEMPTED	PERCENT CORRECT	RT CORRECT	RT ERROR
SLIGHT	9.0	6.0	-6.0	*
MODERATE	7.0	4.0	4.0	*
SEVERE	-5.0	-2.0	-4.0	*

*Not computable
RT = Reaction time

As expected psychomotor behavior was relatively insensitive to the acute effects of the heat stress. Despite the small degree of change (Table 12), accuracy was down in the severely heat stressed group. RTcor and RTerr changed differently from each other, with correct responses being emitted slower while RTerr being emitted quicker due to severe heat stress.

TABLE 12

PERCENT CHANGE IN REACTION TIME TESTS

	NUMBER ATTEMPTED	PERCENT CORRECT	RT CORRECT	RT ERROR
CONTROL	3.0	0.0	-3.0	10.0
SLIGHT	2.0	0.0	-3.0	6.0
MODERATE	0.0	1.0	-1.0	0.0
SEVERE	-2.0	0.0	3.0	-10.0

RT = Reaction time

Mood data were idiosyncratic and varied independently of the stress encountered. Figure 6 presents the activation and mood scores for the three subjects in the severe stress group. As can be seen, some subjects reported changes while others reported no changes. All subjects seemed to be less active and in a worse mood after the severe heat stress condition, but by widely differing amounts.

```
                ACTIVATION
LETHARGIC                        FULL.OF ENERGY
  +----+----+----+-A--+B---+----+----+----+----+
-4+----+----+----+----+A--B+----+----+----+----+4
  +----+----+----+-A--+----+-B--+----+----+----+

                   MOOD
GOOD                                        BAD
  +----+--B-+A---+----+----+----+----+----+----+
-4+----+--BA+----+----+----+----+----+----+----+4
  +-B--+---A+----+----+----+----+----+----+----+

               B (Before)
               A (After)
```

FIGURE 6. Scaled Self-Reports of Mood of Severely Heat-Stressed Subjects. B represents the preflight report, A the postflight report, and N no change.

These results cannot support the position that psychological/psychomotor function varied systematically as a function of the type of CD ensemble worn while flying. However, if the grouping of subjects into the arbitrary classes of slightly, moderately, and severely heat stressed is accepted, then trends emerge which the author believes are systematic and confirm the applicability of laboratory investigations of heat stress to the aviation setting.

The data reported here suggest that slight heat stress increases performance over control levels[21] and that this improvement is eliminated by more severe heat stress.[22] These results are probably conservative due to the intervention of an unavoidable recovery period between exposure and posttesting. Without arguing the significance of changes in performance (number attempted and percent correct) or their operational significance. The effect upon reaction time which terminated in error (RTerr) was clearly anomalous. During one test (target detection), subjects spent a great deal more time than expected working on the solution without being able to find the correct answer. Other tests (e.g., logical reasoning) showed that subjects made errors without working on the problem for as long as expected. In other words, when subjects were severely heat-stressed, they either could not provide the correct answer despite extra effort or could not recognize that additional consideration was necessary. The conclusion that subjects failed to adequately consider the problem at hand is based upon the fact that response latencies were shortened without a concomitant increase in error rates. The possibility that subjects chose not to answer a particular question and in that fashion shortened response latencies could not be ruled out. This result has previously been reported by Colquhoun and Goldman[23]

Self-report of mood varied widely across the severely heat stressed subjects. This lack of consistency between self-report and heat stress is not unusual. When describing their behavior, people follow rules which are more in keeping with their social environment than their internal state.[21] Some people will follow the rule that states that the effect of exposure to heat is to slow response times and reduce performance levels. Others follow the rule that a "can do" attitude is important to maintain regardless of the situation. This type of rule-following results in a dissociation between level of cognitive function and reported mood.

Pilot Performance Factors

Pilot flight performance was assessed by obtaining inflight aircraft status data for three flight profiles, a heading altitude, airspeed and time (HAAT) maneuver, a lateral hover, and a 50-foot out-of-ground-effect hover. These three maneuver exercises were performed in sequence for the 4-hour flight period. The HAAT maneuver consisted of nine trials. At the beginning of a trial, the safety pilot would read a set of parameters to the subject pilot. The subject could then ask for one repeat or proceed to comply with the instructions. Upon compliance the subject said "start" and maintained those parameters until he believed he had reached the prescribed time limit. When he acknowledged the end of the time period, a new trial was begun. An example of the parameters instruction would be: "Heading, one-eight-zero degrees; altitude, nine hundred feet; airspeed, eighty knots; time, twenty seconds." The lateral hover exercise required the subject to bring the aircraft to a stabilized hover and hover laterally around a rectangular runway, keeping the mast of the aircraft over the edge of the runway. At each corner, the subject performed a 450-degree pedal turn before continuing. The exercise ended when the subject completed the rectangular course by returning to the starting point. The 50-foot hover exercise required the subject to hover the aircraft into the wind at a perceived altitude of 50 feet above ground level for a period of 2 minutes.

Dependent variables analyzed for the flying tasks were:

1. Absolute difference between the instructed heading and the observed heading at the start of each trial.

2. Absolute difference between the instructed airspeed and the observed airspeed at the start of each trial.

3. Absolute difference between the instructed time and the observed elapsed time between the start and stop of each trial.

4. The median of heading standard deviations between the start and stop for each trial.

5. The median of airspeed standard deviations between the start and stop of each trial.

6. Elapsed time for acknowledgment of instructions to "start" for trials two, four, and seven.

Variables one through three are measures of accuracy of compliance with instructed parameters at the beginning of a trial. Variables four and five are measures of how well the subject maintained instructed straight and level flight. Variable six is a measure of subject response latency on changing aircraft flight parameters.

Statistical analysis of the data collected to assess these variables provided only one significant comparison differentiating between suits on the basis of performance; that of median heading error at the start of the straight and level flight ($p=.01$). All other comparisons were nonsignificant. Further testing on the median heading error measure revealed that the significance of this difference could not be attributed to any one ensemble and thus was judged to have no practical effect. A review of data of the two subjects who were terminated for exceeding heat safety criteria while in the US ensemble also revealed satisfactory flight performance just prior to termination. In fact, the performance immediately before termination was not distinguishable from flight performance measured earlier in the day.

In review of the results of this exercise, three of the six subjects wearing the US ensemble were moderately heat stressed as judged by heart rates and rectal temperatures and either quit or were terminated by the medical observer. All other subjects flew the prescribed 4-hour mission or were prevented from doing so by aircraft maintenance problems or adverse weather. One of this group was also terminated (UK ensemble) for convergence between chest and core temperature when in fact his core temperature and heart rate showed no sign of heat stress. This event reemphasized to the research team that the judgment call on whether a pilot is becoming stressed involves a pattern recognition of many factors including rectal temperature, heart rate, and skin temperature. As the study progressed, it became obvious that the heart rate changed rapidly as a subject interacted with his environment and that heat stress tended to bias the heart rate upward on a more continuous basis.

Those aviators in this study who were heavier for their height and older than 29 were found to be more susceptible to heat stress than their younger, lighter counterparts. This observation is consistent with comments made by Goldman[24] in reviewing predictors of heat strain and Myhre's[25] findings that those who completed the task of rapid runway repair while wearing CD ensembles in hot weather were younger and in better physical shape as measured by aerobic capacity.

The safety limits of HR > 140 bpm for 15 minutes or rectal temperature > 38.5°C or convergence to within 0.5°C of mean skin temperature and rectal temperature were more conservative than normally used in stress research because of extra concerns for safety in the inflight environment. Their *a priori* selection narrowed the range of possible responses to such an extent that in two of the cases where aviators were terminated for high heart rate it was apparent that they could have flown longer. Thus, none of the aviators, except for the aviator who quit, were severely heat stressed. Nonetheless, consistent patterns represented in these data do suggest that CD ensembles cause increased rectal temperatures and heart rates in aviators during hot weather and that the US ensemble is somewhat more stressful.

Water was given to all aviators on each flight. It is our opinion that the water was generally beneficial and that drinking regularly helps to forestall the inevitable heat stress which these ensembles will impose in a hot environment during flight training or during combat. However, since water was not withheld as a control, there is no clear evidence within this study to substantiate this general feeling. On the other hand, it is quite clear that exposure to high radiant heat loads outside the aircraft combined with the muscular activity of walking to the rest area during the refueling break resulted in an increase in stored energy in many of the subjects. Additional work, such as arming or refueling, would exacerbate this problem and result in susceptible aviators becoming casualties sooner.

Analysis of the cognitive data obtained during this exercise did not support the position that psychological/psychomotor function varies systematically with the type of ensemble worn while flying. Nonetheless, as with the physiological data, information obtained from this work in the field setting does serve to confirm the applicability of previous laboratory research findings of heat stress to the aviation environment.

The finding that there were no practical differences in pilot performance while wearing the ST suit, the US, and UK CD ensembles is significant. In addition to this, it should be noted that performance during normal flight did not serve as a predictor or indicator of heat stress. Although the criteria used for termination were admittedly conservative, subjects experienced heat stress only in the US CD ensemble, and they were able to maintain performance up to the point of reaching termination criterion. This finding is supported by the fact that the US ensemble has the highest CLO (insulation) value of the three ensembles tested.

CONCLUSIONS

Generally, the subjects were not stressed in this study to the extent that they would have been in an operational scenario. The approach was basically conservative to insure the health and safety of the aircrew. However, physiological indicants of stress, particularly heart rate and to some extent core temperature, seem to be sensitive measures of heat loading which may be used as warning signals for aircrew in some operational environments. This is consistent with the basic science literature assessing the effects of heat stress on physiological function. Cognitive and aviation performance on the other hand do not provide sufficient lead time before the onset of physiological systemic heat load to be very effective as alerting mechanisms.

Substantial research is still required assessing performance, cognitive and physiological factors under heat stress and during sustained operations to effectively assess the full predictive validity of these measures. Further work must be directed at obtaining data under conditions where the subject is placed under a more realistic and extensive work/exercise load in order to predict aircrew staying power in the chemically contaminated battlefield environment.

REFERENCES

1. Department of the Army. Operations (FM 100-5). Washington, DC: Headquarters, Department of the Army, 1982.

2. Alexander M. Haig, Jr. Chemical Warfare in Southeast Asia and Afghanistan (Special Report No. 98). Washington, DC: United States Department of State, March 1982.

3. Garrett, P. F., Jr. Aircrew chemical/biological (CB) protection (Letter Report No. 264). Aberdeen Proving Ground, MD: US Army Human Engineering Laboratory, 1979.

4. Clanahan, H. L. Comparative test chemical/biological (CB) disposable overgarment (OT&E) Final Report). Eglin Air Force Base, FL: Tactical Air Warfare Center, 1974.

5. Chestnut, J. L. Evaluation of M-24 protective assembly (gas) for use by aerial spray crews (TAC 68-1-7). Eglin Air Force Base, FL: USAF Special Operations Center, 1969.

6. Dize, J. H. and Steyer, W. C. Concept evaluation program of the United Kingdom aircrew, NBC defense assembly (9-CEP-054). Fort Rucker, AL: US Army Aviation Development Test Activity, 1979.

7. Bruce E. Hamilton, Dennis Folds, and Ronald R. Simmons. Performance impact of current United States and United Kingdom aircrew chemical defense ensembles (USAARL Report No. 82-9). Fort Rucker, AL: US Army Aeromedical Research Laboratory, September 1982.

8. Francis S. Knox, III, Gerald A. Nagel, Bruce E. Hamilton, Raul P. Olazabal, and Kent A. Kimball. Physiological impact of wearing aircrew chemical defense protective ensembles while flying the UH-1H in hot weather (USAARL Report No. 83-4). Fort Rucker, AL: US Army Aeromedical Research Laboratory, October 1982.

9. Bruce E. Hamilton, Ronald R. Simmons, and Kent A. Kimball. Psychological effects of chemical defense ensemble imposed heat stress on Army aviators (USAARL Report No. 83-6). Fort Rucker, AL: US Army Aeromedical Research Laboratory, November 1982.

10. Livingston, S. D., Rud, C. J., Brooks, C. J., Bowen, C. Comparative heat stress measurements flying with and without a CW ensemble. San Antonio, TX: Preprint Aerospace Medical Association Annual Meeting, 4-7 May 1981.

11. Pandolf, K. B., and Goldman, R. F. Convergence of skin and rectal temperatures as a criterion for heat tolerance. Aviation, Space, and Environmental Medicine. 49(9):1095-1101. 1978.

12. Huffman, H. W., Hofmann, M. A., and Sleeter, M. R. Helicopter in-flight monitoring system (USAARL Report No. 72-11). Fort Rucker, AL: US Army Aeromedical Research Laboratory, 1972.

13. Pearson, R. G., and Byars, G. E., Jr. The development and validation of a checklist for measuring subjective fatigue (Report 56-115). Randolph AFB, TX: USAF School of Aviation Medicine, 1956.

14. Haslam, D. R. The military performance of soldiers in continuous operations: exercise "Early Call" I and II. In: Johnson, L. C., Tepas, D. I., Colquhoun, W. P., and Colligan, M. J. (eds.), The twenty-four hour workday: proceedings of a symposium on variations in work-sleep schedules. Cincinnati: US Department of Health and Human Services (DHHSCNOISH) Publication No. 81-127, 1981.

15. Folkard, S., Knauth, P., Monk, T. H., and Ruten, F. J. The effect of memory load on the ciradian variation in performance efficiency under a rapidly rotating shift system. Ergonomics. 19:479-488. 1976.

16. Baddeley, A. D. A 3-minute reasoning test based on grammatical transformation. Psychonomic Science. 10:341-342. 1968.

17. Weaver, R. A. The circadian system of man. New York-Heidelberg-Berlin: Springer Verlag, 1979.

18. Wilkinson, R. T., and Hougton, D. Portable four choice reaction time test with magnetic tape memory. Behavior Research, Methods and Instrumentation. 7:441-446. 1975.

19. Klein, K. E., and Wegmann, H. M. Significance of ciradian rhythms in aerospace operations. London: Technical Editing and Reproduction Ltd. NATO AGARDograph No. 247, 1980.

20. Edwards, A. L. Experimental design in psychological research. New York: Holt, Rinehart and Winston, Inc., 1960.

21. Poulton, E. C. Arousing environmental stresses can improve performance, whatever people say. Aviation, Space, and Environmental Medicine. 47(11):1193-1204. 1976.

22. Epstein, Y., Keren, G., Moisseiev, J., Gasko, O., and Yachin, S. Psychomotor deterioration during exposure to heat. Aviation, Space, and Environmental Medicine. 51(6):607-610. 1980.

23. Colquhoun, W. P. and Goldman, R. F. Vigilance under induced hyperthermia. Ergonomics. 15(6):621-632. 1972.

24. Goldman, R. F. Prediction of heat strain, Rev ed (USARIEM Report P-60/79, 1979.

25. Myhre, L. G. Field study observations of physiological responses of USAF ground crew performing rapid runway repair activities. Bal Harbour, FL: Aerospace Medical Association Annual Meeting, 10-13 May 1982.

DISCUSSION

COL D R PRICE (US)

Would you please confirm that the limit to the pulse rate at which you terminated a flight was a rate of 140 beats per minute sustained for 15 minutes? Would you also comment please on the flights which were terminated because the skin temperature reached within 0.5°C of the core temperature?

AUTHOR'S REPLY

Yes, you are correct with respect to the points at which we terminated the flights. In retrospect we were unduly cautious. We certainly need not have terminated the flight in the UK ensemble in which skin temperature reached 0.5°C of the core temperature.

E E HOWARD (UK)

You reported that 3 subjects were withdrawn when wearing the US equipment. Were these 3 subjects also tested wearing the UK suit? If so, were there differences attributable to the equipment?

AUTHOR'S REPLY

The experiment was designed as a cross-over study and the same subjects wore both the US and the UK ensemble in turn. We found that the physiological changes were less severe with the UK ensemble. The subjective reports also showed that the UK ensemble was much preferred. The subjects tolerated the fairly hot environment far better when using the UK ensemble than when they wore the US equipment.

DR L C BOER (NL)

You showed cognitive performance on one of your slides but did not present any results. Were the results not worth showing?

AUTHOR'S REPLY

As I stated during the presentation, analysis of the performance assessment battery (psychological and psychomotor tasks) data failed to differentiate between the test conditions, i.e. ensemble, etc. However, when subjects were grouped according to the level of heat stress regardless of the ensemble worn, trends similar to those reported in the classic heat stress literature were evident; but they were not statistically significant.

DR A RIECK (GE)

I would like to propose that you should compare the pulse rate responses in terms of relative values, i.e. percentage of the individual starting value. This approach would eliminate the effects of variations in the starting pulse rates.